Received Date : 20-Jan-2013
Revised Date : 17-May-2013
Accepted Date : 07-Jun-2013
Article type : Original Article
Manuscript Category: Prevention & Epidemiology (PRE)

Evaluation of *PAX3* genetic variants and nevus number

Zighereda Ogbah[1], Celia Badenas[2,3], Mark Harland[4], Joan Anton Puig-Butille[2,3], Fay Elliot[4], Nuria Bonifaci[5], Elisabet Guino[6], Julie Randerson-Moor[4], May Chan[4], Mark M Iles[4], Daniel Glass[7,8], Andrew A Brown[7], Cristina Carrera[1,3], Isabel Kolm[1], Veronique Bataille[7] , Timothy D Spector[7], Josep Malvehy[1,3], Julia Newton- Bishop[4], Miquel Angel Pujana[5], Tim Bishop[4], Susana Puig[1,3]

1 Melanoma Unit, Department of Dermatology Hospital Clínic de Barcelona, IDIBAPS, Barcelona University, Spain

2 Melanoma Unit, Biochemistry and Molecular Genetics Service, Hospital Clínic de Barcelona, IDIBAPS, Barcelona University, Spain

3 Centre of Biomedical Research on Rare Diseases (CIBERER), ISCIII, Barcelona, Spain

4 Division of Epidemiology and Biostatistics, Leeds Institute of Molecular Medicine (LIMM), University of Leeds, Leeds, United Kingdom

5 Breast Cancer and Systems Biology Unit, Translational Research Laboratory, Catalan Institute of Oncology, Bellvitge Biomedical Research Institute (IDIBELL), L'Hospitalet, Barcelona, Spain

6 Biomarkers and Susceptibility Unit, Catalan Institute of Oncology, IDIBELL, L'Hospitalet, Barcelona, Spain

7 Department of Twin Research & Genetic Epidemiology, Kings College London, St. Thomas' Hospital Campus, London, UK

8 *Department of Dermatology*, *The North West London Hospitals NHS Trust, London, UK*

Correspondence:
Zighereda Ogbah
Melanoma Unit, Dermatology Department, Hospital Clinic Barcelona
Villarroel 170, 08036 Barcelona. Spain; zigheog@gmail.com

Susana Puig, MD Consultant
Melanoma Unit, Dermatology Department, Hospital Clinic Barcelona
Villarroel 170, 08036 Barcelona. Spain; susipuig@gmail.com or spuig@clinic.ub.es

Tel.: +34 93 2275400 ext 2893. Fax: +34 93 2275438

Summary

The presence of a high nevus number is the strongest phenotypic predictor of melanoma risk. Here, we describe the results of a three-stage study directed at identifying risk variants for the high nevus phenotype. At the first stage, 263 melanoma cases from Barcelona were genotyped for 223 single nucleotide polymorphisms (SNPs) in 39 candidate genes. Seven SNPs in the *PAX3* gene were found to be significantly associated with nevus number under the additive model. Next, the associations for seven *PAX3* variants were evaluated in 1,217 melanoma cases and 475 controls from Leeds; and in 3,054 healthy twins from TwinsUK. Associations with high nevus number were detected for rs6754024 (P values < 0.01) in the Barcelona and Leeds datasets and for rs2855268 (P values < 0.01) in the Barcelona and the TwinsUK sets. Associations (P values < 0.001) in the opposite direction were detected for rs7600206 and rs12995399 in the Barcelona and TwinsUK sets. This study suggests that SNPs in *PAX3* are associated with nevus number, providing support for *PAX3* as a candidate nevus gene. Further studies are needed to examine the role of *PAX3* in melanoma susceptibility.

Significance

Common acquired nevi are melanocytes in proliferation and high nevus number represents the most important phenotypic melanoma risk factor. Therefore, understanding the genetic bases of nevi is crucial to better understand the etiology of melanoma and to develop better predictions of melanoma risk. Genome-wide association studies on nevus number have identified a number of loci that explain only a small proportion of the risk, suggesting that additional loci remain to be identified. Through a candidate gene study, using cohorts from Spain and UK, our results suggest that genetic variants at the *PAX3* locus are associated with nevus number.

KEY WORDS: nevus, melanoma, PAX3, SNP, susceptibility

Running Title: *PAX3* genetic variants and nevus number

Introduction

Cutaneous melanoma is a potentially fatal type of skin cancer with increasing incidence worldwide (Leiter and Garbe, 2008). The strongest phenotypic risk factor for melanoma development is the presence of high nevus number (Bataille et al., 1996; Briollais et al., 2000; Gandini et al., 2005; Chang et al., 2009). Common acquired nevi are benign melanocytic proliferations that usually appear in early childhood followed by a decline in number after middle adulthood. Both genetic and environmental factors are involved in either nevus-genesis or melanoma-genesis. Nevertheless, the association between sunlight

exposure, melanoma and nevus development is relatively complex. Twin studies have provided strong evidence that nevus number is predominantly genetically determined (Easton et al., 1991; Zhu G.et al., 1999; Bataille et al., 2000; Wachsmuth et al., 2001) with a smaller effect of sun exposure (Wachsmuth et al., 2005). Familial aggregation of nevus number and melanoma has also been described (Bahuau et al., 1997; Briollais et al., 2000; Florell et al., 2004). However, mutations in known melanoma susceptibility genes do not always co-segregate with nevus number (Puig et al., 1997; Bishop et al., 2000) suggesting that the mode of inheritance of nevus number is complex (Bataille et al., 2000). Linkage studies in healthy twins identified MTAP, at a locus associated with nevus near *CDKN2A* in addition to other loci, showing that nevus number is likely to be a complex trait with many genes involved (Falchi et al., 2006, Zhu et al., 2007). Genome-wide association studies (GWASs) have identified variants in *MTAP* on 9p21, *PLA2G6* on 22q13.1 (Falchi et al., 2009), *NID1*on 1q34 (Nan et al., 2010) and *IRF4* on 6p25 (Duffy et al., 2011) in association with nevus number.

Whilst variants involved in nevus are unlikely to fully explain melanoma risk and vice versa, more knowledge on the genetic basis of nevus development is needed, as it constitutes the strongest risk predictor for sporadic melanoma (Bataille et al., 1996; Briollais et al., 2000; Gandini et al, 2005; Chang et al, 2009). Here, we approach this question through a candidate gene study with replication in two different populations.

Results

Association study between nevus number and genetic variants

Barcelona melanoma cases

Of the 223 SNPs selected from the 39 candidate genes, 13 showed a call rate < 90% and, therefore, were excluded from subsequent analysis. Duplicate concordance rate was >90%. Next, under the additive model, 21 SNPs showed significant associations with nevus number at a false discovery rate (FDR) < 5% (Supplementary Table 1.2). Seven of these SNPs were located at the *PAX3* locus (Supplementary Figure 1). . Given these results and the known role of PAX3 in promoting melanoma progression (Scholl et al., 2001; Plummer et al., 2008), subsequent studies were focused on this locus. Of the seven *PAX3* genetic variants, rs6754024 (OR 2.00; 95% CI (1.20-3.34); $P_{additive}$=0.0066), rs10180903 (OR 2.54; 95% CI (1.51-4.29); $P_{additive}$=0.0003) and rs1013262 (OR 2.02; 95% CI (1.39-2.93); $P_{additive}$= 0.0001) were associated with high nevus number, while rs7600206 (OR 0.38; 95% CI (0.22-0.66); $P_{additive}$=0.0002), rs12995399 (OR 0.46; 95% CI (0.29-0.74); $P_{additive}$=0.0008), rs2855268 (OR 0.46; 95% CI (0.26-0.84); $P_{additive}$=0.008) and rs1978859 (OR 0.47; 95% CI (0.30-0.75); $P_{additive}$=0.0009) were associated with low nevus number (Table 1).

Leeds case-control study

To replicate the results from the Barcelona study, 1,217 melanoma cases and 475 controls, collected from the Leeds case control study (Randerson-Moor et al., 2009), were genotyped for the seven *PAX3* SNPs described above. The SNP rs1978859 failed to genotype and rs7600206 did not show HWE in controls. The concordance rate between duplicates was 100%.

Association between SNPs and nevus number were determined for cases and controls combined and then separately. In a combined analysis, an association was found with high nevus number for rs6754024 under a recessive model, adjusting for age, gender and melanoma status (OR 3.43; 95% CI (1.55-7.62); $P_{recessive}$=0.0037). When the analysis was stratified by cases and controls, the positive association with high nevus number came only from the melanoma cases (OR 3.57; 95% CI (1.50-8.52); $P_{recessive}$=0.0053) (Table 1), as in the BCN study, whereas among Leeds controls, no association was observed (OR 2.84; 95% CI (0.33-24.58); $P_{recessive}$=0.399 (Supplementary Table 2.1). For SNP rs6754024, GT heterozygote cases with less than 50 nevi are more frequent (27%) in Leeds than in BCN (20%) while GT heterozygote cases with >100 nevi are more frequent in BCN (29%) than in Leeds (23.1%) (Tables 1 and 2). In the Leeds dataset, the low number of heterozygotes in cases with >100 nevi results in decreasing OR for heterozygotes (0.83) and increasing OR for rare homozygotes (3.39). Due to this, the T allele for SNP rs6754024 showed a significant recessive effect on nevus number among Leeds cases (Tables 1 and 2). HWE for SNP rs6754024 was tested among Leeds controls (P=0.545) and cases (P=0.395). The difference in heterozygote numbers among cases could probably result from the difference in design between BCN and Leeds studies, rather than from genotyping quality. The BCN cohort differs from Leeds in that it is not population based and is enriched for high risk patients with >100 nevi and multiple melanomas.

To exclude that the rs6754024 association with nevus number in Leeds was due to different phenotypes, the analyses were repeated including only Leeds cases with 50-100 nevi and considering nevus number as a continuous variable. These analyses confirmed that rs6754024 was the only SNP associated with high nevus number among Leeds cases and that only rs6754024 TT cases ($P_{recessive}$=0.006) showed an increased risk of high nevus number when compared with GG and GT cases (Supplementary Table 2.2).

A joint analysis of the BCN and Leeds cases after adjustment for age, sex and centres confirmed the association between rs6754024 and risk of high nevus number under a recessive model (Table 2). The stronger association of rs6754024 with nevus number in the joint analysis was primarily driven by the Leeds cases where rs6754024 seems to act in a recessive manner (Table 2). In BCN and Leeds cases, the rs6754024 minor allele T was associated with high nevus number (BCN P=0.0066; Leeds P=0.0053) with similar frequencies (BCN= 0.15; Leeds=0.16). Genotype frequency comparison showed no heterogeneity between cases from BCN and Leeds (P= 0.8).

Association between the SNP rs6754024 and potential confounding factors, such as skin type and multiple primary melanomas, was assessed for BCN cases and Leeds cases. No associations were observed between rs6754024 and skin type (BCN, OR 0.77; 95% CI (0.45-1.31); $P_{additive}$=0.329; Leeds, OR 1.15; 95% CI (0.57-2.32); $P_{recessive}$=0.703) or number of primary melanomas (BCN, OR 1.16; 95% CI (0.71-1.91); $P_{additive}$=0.554; Leeds, OR 1.96; 95% CI (0.58-6.62); $P_{recessive}$=0.316).

TwinsUK study

We then attempted to replicate the association with nevus number for the seven *PAX3* SNPs, found in BCN, in a healthy twin cohort from TwinsUK. Since only two SNPs within PAX3 (rs1013262 and rs12995399) were directly genotyped, imputed SNPs were used to evaluate associations.

Four out of seven SNPs were found associated with nevus number. Using either total nevus number or log nevus number (Table3a), the strongest associations were for rs7600206 ($P_{additive}$=0.0001; or $P_{additive}$=0.0005), rs12995399 ($P_{additive}$=0.0005; or $P_{additive}$=0.0032) and rs2855268 ($P_{additive}$=0.0503; or $P_{additive}$=0.0038). The association for rs1013262 was also significant for total nevus number ($P_{additive}$=0.0179) but after log nevus number analysis, it was no longer significant ($P_{additive}$=0.347) (Table3a).

The SNP rs2855268 was associated with decreased risk of high nevus number, as in the BCN study, whereas rs7600206 and rs12995399 were associated with increased risk of high nevus number, the opposite to that observed in the BCN study.

After taking family structure into account for the twin relatedness, log nevus number analysis still showed significant associations for SNPs rs7600206 ($P_{additive}$=0.0020), rs12995399 ($P_{additive}$=0.0163) and rs2855268 ($P_{additive}$=0.0057) (Table3a).

In the proxy analysis, an association with nevus number was found for rs10932952 ($P_{additive}$=0.0013) and rs7588554 ($P_{additive}$=0.0276) (Table 3b), which are in complete LD (r^2=1) with candidate SNPs rs12995399 and rs1013262, respectively (Table 3c). Singletons (half twin population) analysis confirmed the association between SNPs rs7600206 ($P_{additive}$=0.0028) and rs12995399 ($P_{additive}$=0.0197) and nevus number (Supplementary Table 2.3).

Association between *PAX3* SNPs and skin type, as a potential confounding factor, was not significant (rs7600206 $P_{additive}$=0.296; rs12995399 $P_{additive}$=0.236; rs2855268 $P_{additive}$=0.405).

TwinsUK includes only females, therefore, to complete the study, we examined whether *PAX3* SNPs were associated with gender in BCN and Leeds cases. No gender-specific association was found in stratified analysis (data not shown).

Functional SNP Prediction Analysis

The potential functional relevance of the four *PAX3* SNPs associated with nevus number in at least two of the three studies (rs6754024, rs7600206, rs12995399, and rs2855268) was evaluated, examining data from expression quantitative loci (eQTL) (GenVAR http://www.sanger.ac.uk). The expression data is from normal skin biopsies from a large subset of the twins in the TwinsUK cohort from the Muther Consortium (www.muther.ac.uk). In this analysis, HapMap SNPs in linkage disequilibrium (r^2>0.5 in European population) with the four associated variants were also examined. While rs2855268 was not present in the eQTL dataset, lower *PAX3* expression was observed for skin samples with the rs6754024 AA genotype (*P*= 0.002 adjusted by permutation test) (Supplementary Figure 2A). Other SNPs in LD were found associated with *PAX3* expression differences in skin samples: rs6436309 (in LD with rs6754024, r^2=0.53), and rs1549773 and rs13410020 (in LD with rs2855268, r^2=1 and r^2=0.68, respectively). These revealed associations with lower expression for the minor genotypes (rs6436309 *P*=0.002; rs1549773 *P*=0.006; rs13410020 *P*=0.009, adjusted by permutation test) (Supplementary Figure 2B, 2C and 2D). No associations were found when other tissue types, as fat tissue or lymphoblastoid cell lines, were considered, which suggests a specific effect in for the causal variant influencing nevus biology.

Further functional analyses of the associated SNPs predicted that rs6754024 alters *PAX3* transcription factor binding activity (NIEHS http://egp.gs.washington.edu) whereas rs2855268 may have a splicing silencer activity (ACESCAN2 http://genes.mit.edu/acescan2/index.html; ESRsearch http://esrsearch.tau.ac.il). Together, these observations provide a mechanistic hypothesis for the risk of high nevus number associated with altered *PAX3* function.

Discussion

Low-penetrance genes are involved in nevogenesis and, in turn, in melanoma susceptibility. Our study evaluated the association of *PAX3* polymorphisms with nevus number. This gene encodes for a transcription factor that plays a central role in melanocyte development and malignant melanoma (Lang et al., 2005). PAX3 is expressed in nevi and melanoma (Scholl et al., 2001, Plummer et al., 2008) and is also considered as a marker of melanoma staging and circulating melanoma cells (Koyanagi et al., 2005, Takeuchi et al., 2004). Moreover, MITF, a direct downstream target of PAX3, confers a predisposition to melanoma by the identification of a germline missense variant in familial and sporadic melanoma, (Yokoyama et al., 2011).

A total of seven *PAX3* variants were found to be associated with high nevus number after adjusting for multiple tests in an enriched case-only population comprising 263 cases from BCN. To validate these results, the variants were evaluated in two independent replication sets from the UK. The SNP rs6754024 was associated with risk of high nevus number in the melanoma cases from BCN and Leeds. The rs6754024 seems to act in a log-additive manner in the BCN cases and in a recessive manner in the Leeds cases. In both case groups, the rs6754024 minor allele T was associated with high nevus number with similar

frequency. We observed that the difference in the models for rs6754024 between the two cohorts is due to the low number of heterozygotes in Leeds cases. Probably, this is due to the different study design of the two populations, the BCN set being enriched for the presence of high risk cases with more than 100 nevi and multiple primary melanomas whereas the Leeds cohort was population based.

The SNP rs6754024 was found associated with nevus number in melanoma cases from BCN and Leeds suggesting a possible involvement of *PAX3* in melanoma susceptibility. The association with nevus number for SNP rs6754024 was independent of potential confounding factors such as skin type, number of primary melanomas and gender status in both cohorts. The rs6754024 was not associated with nevus number in controls from Leeds nor among TwinsUK, a cohort that included a larger number of healthy individuals. This could be related to the fact that, in Leeds, nevus number was higher among melanoma cases compared to controls. The positive association of SNP rs6754024 only among melanoma cases suggests that PAX3 is involved in melanoma susceptibility.

By using public web tools and eQTL data sets from the Muther dataset derived from the same twin cohort from TwinsUK, a functional effect of SNP rs6754024 was predicted. In fact, a decreased *PAX3* expression in skin from TwinsUK was correlated with the presence of the minor allele of SNP rs6754024. Since this SNP was associated with high nevus number among melanoma cases, it can be hypothesized that the risk of having high nevus number and/or melanoma is correlated with increased *PAX3* expression, as shown by previous reports in nevi and melanoma tumours (Scholl et al., 2001, Plummer et al., 2008).

In the healthy TwinsUK cohort, the association with nevus number was confirmed for three out of seven *PAX3* SNPs: rs2855268, rs7600206 and rs12995399. These associations were confirmed to be independent of skin type, which is a confounding factor for nevi. The relationship with nevus number in the BCN cases and in the healthy twins suggests a potential involvement of PAX3 function in nevus susceptibility.

In TwinsUK, rs2855268 was associated with decreased risk of high nevus number as found in the BCN study. The rs2855268 variant is located 19 bp upstream of exon 8 and was predicted to have splicing silencer activity. *PAX3* produces up to nine different isoforms in man, which result from alternative splicing (Medic and Ziman, 2009). In mice myoblasts, alternative splicing of *Pax3*, including exons 7, 8 and 9, produces an alternatively spliced transcriptionally inactive form, which is able to inhibit the transcriptional activity of the PAX3 protein (Pritchard et al., 2002). In melanoma, the same region has shown to be involved in the generation of abundant *PAX3* transcripts in human melanocytes (School et al., 2001). Moreover two SNPs, rs1549773 and rs13410020, in LD with rs2855268 have been shown to be associated with low *PAX3* expression level in skin samples from TwinsUK. It could therefore be speculated that rs2855268 might be involved in *PAX3* expression regulation.

In TwinsUK, rs12995399 and rs7600206 were found to be associated with increased risk of high nevus number, the opposite direction of effect found in the BCN cases where they protect from high nevus number. The opposite effect of the SNPs observed in these two cohorts is not entirely biologically inconceivable as it may be related to population differences. In addition, the opposite direction could be determined by the presence of different haplotypes with one or more causal variants, which might be implicated in biological process with different effects in melanoma cases and healthy controls. Different functions have been described for *PAX3* in melanocytes and melanoma (Medic and Ziman, 2010). *PAX3* is a recognized key regulator of melanocytes development during embryogenesis (Blake and Ziman, 2005, Kubic et al., 2008) but in adult melanocytes the exact role of *PAX3* is still unclear (Medic et al., 2011). PAX3 re-expression in adult melanocytes has not been shown to promote melanoma progression as previously stated (Medic and Ziman , 2009) as recent reports have shown that PAX3 expression is observed in melanocytes in normal skin as well as in benign nevi (He et al., 2010, Medic and Ziman, 2010). PAX3 expression, specifically its up-regulation, in melanocytes in normal skin, has been shown to be the result of UV-induced loss of TGF-beta signalling from keratinocytes, suggesting a key role for PAX3 in UV-pigmentation response (Yang G et al., 2008). PAX3 regulates the proliferation and survival in both melanoma cells and melanocytes from normal skin whereas it seems to regulate migration in melanocytes from nevi and melanoma but not in melanocytes from normal skin (Medic and Ziman, 2010). In melanoma, PAX3 has been shown to promote melanoma invasiveness downstream from PI3K signaling, by its novel capacity to bind to the Brn-2 promoter and to activate Brn-2 expression (Bonvin et al. 2012). Interestingly, it has been suggested that PAX3 in melanocytes from normal skin might regulate malignant transformation of a sub-population of melanocytes expressing PAX3 and displaying a malignant phenotype, characterized by lack of apoptosis, non-differentiation and proliferation (Medic and Ziman, 2010).

Searching for causal variant candidates, only rs6754024 was found completely in LD (r^2=1) with a non intronic variant, rs12623857 (p.Gly43), a synonymous SNP, previously associated with phenotypic features related to melanoma risk, such as dark skin color and light hair color, in a Spanish case-control study (Fernandez et al., 2007). Although a synonymous SNP in a coding region is considered to be a coding SNP that does not change the protein sequence, studies have demonstrated that synonymous polymorphisms can affect messenger RNA splicing, stability, and structure as well as the rate of protein folding (Sauna and Kimchi-Sarfaty, 2011). The *PAX3* gene contains two domains to bind cognate DNA sequences of target genes, a paired domain and a homeodomain (Chalepakis et al., 1994). Mutations in paired domain in melanocytes affect PAX3 transcription activity (Corry and Underhill, 2005). The paired domain is encoded by exons 1-4 and, therefore, SNPs lying within this genomic region, such as rs6754024 and rs12623857, located in intron 1 and exon 2 respectively, could be of functional importance.

Nevus number is a risk factor tightly linked to melanoma. The strength of this study is that we showed a relationship between PAX3 and nevus number in three independent populations, confirming the value of

studying nevi genetics in the search for new melanoma genes. Our study was limited by the different study designs for the three populations which made the interpretation of the results difficult: the association analyses were principally strictly evaluated by trend models in the BCN enriched cases and TwinsUK. In the Leeds cases, analyses were evaluated for multiple genetic models because prior studies have not specifically implicated one inheritance model over another. SNP function predictions are interesting but show some flaws: first, it is not clear why in GeneVar analysis all significant results were found in Twins group 2 and not in group 1, Twins being separated from the same pair. Secondly, PAX3 presence or absence in normal melanocytes is still under discussion (He et al., 2010, Medic and Ziman, 2010).
Replication of our findings in additional populations is therefore essential to corroborate the results shown here in nevus.

To our knowledge this is the first report of an association between *PAX3* and nevus number, the strongest phenotypic risk factor for melanoma.

From seven *PAX3* SNPs associated with nevus number in the BCN cases, four SNPs were replicated in two cohorts from the UK. In the first instance, the association between rs6754024 and nevus number among melanoma cases and not controls from Leeds suggested that *PAX3* is also involved in melanoma susceptibility. To confirm that the association seen in melanoma cases in Leeds is driven by high nevus number, these results were revalidated in a third cohort, TwinsUK, including a large number of healthy individuals. In this set, 3 SNPs, rs2855268, rs7600206 and rs12995399, were associated with high nevus number, suggesting that *PAX3* is involved in nevogenesis. Functional analyses predict that SNPs rs6754024 and rs2855268 could have functional relevance, directly or indirectly, by affecting PAX3 expression and/or function. Clearly, all these findings suggest a complex role of PAX3 in nevogenesis and melanoma susceptibility.

In conclusion, our findings suggest that *PAX3* genetic variants are associated with nevus number. Further studies are warranted to support *PAX3* as a nevus/melanoma predisposing gene and to improve our understanding of the role of PAX3 in melanoma susceptibility.

Methods

Barcelona melanoma cases

Stage 1 included 263 melanoma patients recruited between 2004-2007 from the Melanoma Unit at the Hospital Clinic of Barcelona, which is a tertiary referral center for melanoma in Catalonia, Spain. Cases were selected according to the total nevus number and classified into two groups: 150 cases with a low nevus number (< 50 nevi) and 113 cases with a high nevus number (> 100 nevi). We selected two groups of patients where there would be no possible overlap between them, as <50 nevi was closer to the median nevus number in our population (Aguilera et al., 2009) and >100 nevi being a well-documented risk factor

for melanoma (Gandini et al., 2005). Both single and multiple primary melanoma were included in the study. Cases with >100 nevi were more likely to have multiple primaries than cases with < 50 nevi (55% *versus* 20%).The mean age at diagnosis of melanoma was 46 years old (range 14-80), but cases with multiple nevi had younger age of onset (41.5 years old) compared to cases with fewer nevi (49 years old) ($P < 0.001$). Total body nevus number was examined by a dermatologist and trained nurses. Nevi ≥2 mm in diameter were counted.

Data on gender, age at diagnosis of melanoma, nevus number, skin type according to Fitzpatrick classification, eye and hair color were collected through a nurse or doctor administered questionnaire. The clinical characteristics of the BCN sample set are described in Supplementary Table 2.4.

All participants gave written informed consent and the study was approved by the institutional ethics committee.

Leeds case-control

To replicate stage 1, a case-control cohort from Leeds (UK) was studied, which included 1,254 cases of melanoma and 483 controls from a geographically defined area of Yorkshire and recruited between 2000 and 2005 at the Genetics Epidemiology Division of St James's University Hospital. The Leeds case-control population was described in a previous study (Randerson-Moor et al., 2009). The mean age at diagnosis of melanoma was 53 years old (range 18-86). Controls (mean age at interview 60 years old; range 22-89) were identified by the cases' family general practitioners as not having any type of cancer and no family history of melanoma.

To apply the same criteria used in the BCN set, the nevus count was categorized into three groups: <50, >100 and between 50 and 100 (50-100) nevi. Among controls, 88% ($n = 425$) individuals had < 50 nevi, 9% ($n = 45$) 50-100 nevi, and 2% ($n = 11$) > 100 nevi. Among cases, 58% ($n = 562$) had < 50, 22% ($n = 211$) 50-100 nevi, and 20% ($n = 191$) had > 100 nevi—therefore, consistent with previous reports, nevus number was a strong predictor of melanoma risk ($P<0.0001$) in this population. A trained nurse interviewed all cases and controls and performed a standardized examination of skin nevi. Nevi were examined by body sector; nevi greater than 2 mm in diameter were counted. Clinical characteristics of patients such as gender, nevus number, age at diagnosis of melanoma, skin type according to Fitzpatrick criteria, eye and hair color were collected as shown in Supplementary Table 2.5. Skin type was strongly associated with melanoma risk ($P<0.0001$). All participants gave informed consent and regional ethical committee approval was obtained.

TwinsUK

A third population was included to confirm the results found in the BCN study. The St Thomas' United Kingdom (UK) adult twin registry (hereafter TwinsUK) cohort was unselected for any disease and is representative of the general UK population.

The TwinsUK nevus data come from 3054 subjects, of which 1494 are singletons. The median age was 47 (range 18.4- 79.5). The median number of nevi was 33 (range 1- 299). The twins were interviewed at St Thomas Hospital in London. All participants provided written informed consent, and the study was approved by the local ethics committee. Examination was performed by trained research nurses following a standardized and reproducible nevus number protocol (Bataille V, 2000). The total body nevus number (excluding the genital area, breasts, and posterior scalp) was defined as the sum of all nevi > 2 mm in diameter. The nevus number protocol in the TwinsUK study was identical to that of the Leeds study.

Genes and polymorphisms selection

A total of 223 SNPs (170 haplotype tagging (htSNPs) and 53 predicted to be functionally relevant) were selected for 39 candidate genes that included known oncogenes for melanoma carcinogenesis as well as genes involved in DNA repair, cell cycle regulation, pigmentation, and melanocyte biology (Supplementary Table 1.1). All SNPs were selected as having minor allele frequencies (MAF) $\geq$ 0.05 in HapMap (www.hapmap.org) European individuals. The htSNPs were selected using data from the HapMap project and the Tagger tool (de Bakker et al, 2005) with linkage disequilibrium $r^2 > 0.5$. Potential functionally relevant SNPs were selected using the PupaSNPs tool (Conde et al. 2004).

By Genevar (http://www.sanger.ac.uk) (Yang et al., 2010) SNPs influencing expression levels of the correspondent gene were analyzed. For this we used expression quantitative trait loci (eQTL) data generated from the Muther Consortium including healthy female Twins from the UK (Nica et al., 2011). Twin pairs where separated into two unrelated groups, thereby performing two independent eQTL analyses. The eQTL analysis was performed in lymphoblastoid cell lines (n=166), skin biopsies (n=160) and fat tissue (n=160). Genevar provides Spearman's rank correlation coefficient (rho) estimates for the strength of relationships between alleles and gene expression intensities for each study group. Furthermore, to test the significance of the relationship, adjusted non-parametric permutation P-values are also provided. The NIEHS (http://egp.gs.washington.edu), ACESCAN2 (http://genes.mit.edu/acescan2/index.html) and ESRsearch (http://esrsearch.tau.ac.il) web tools were used for SNP function prediction.

Genotyping

Genomic DNA was isolated from peripheral blood samples using the salting out procedure. For Barcelona samples, genotyping was carried out using SNPstream Genotyping System (Beckman & Coulter Inc. Fullerton, CA), according to the manufacturer's protocols. The genotyping assay was performed at the

Leeds Institute Molecular Medicine (LIMM) genotyping facilities (Leeds, UK). For the Leeds samples, SNPs were genotyped using the Illumina's VeraCode technology (Illumina, Inc., San Diego, CA, USA). In addition to controls provided by the company, 21 samples corresponding to seven HapMap reference trios were used as internal controls of the genotyping process and for the clustering process. For the Barcelona samples, the genotyping assay was performed at the CeGen genotyping facilities, in the Barcelona Node (Centro Nacional de Genotipado, Genoma España). The total number of duplicates genotyped, intra and inter-assays, were 35/263 for BCN samples and 264/1737 for Leeds samples. For genotyping assays, samples and SNPs with call rates below 90% were excluded from the analysis and failed genotypes were not repeated.

The TwinsUK sample was genotyped with the use of the Infinium 610K assay (Illumina). Population substructure was detected and removed as described previously (Richards JB 2008). Genotypes were cleaned before analysis by removing individuals or SNPs not satisfying the quality control criteria. For samples, exclusion criteria were: (i) sample call rate at least <98%, (ii) heterozygosity across all SNPs ≥2 s.d. from the sample mean; (iii) evidence of non-European ancestry as assessed by PCA comparison with HapMap3 populations; (iv) observed pairwise IBD probabilities suggestive of sample identity errors; (v) concordance at duplicate samples <1%. At the SNP level, exclusion criteria were (i) Hardy-Weinberg p-value<10^{-6}, assessed in a set of unrelated samples; (ii) MAF<1%, assessed in a set of unrelated samples; (iii) SNP call rate <97% (SNPs with MAF≥5%) or < 99% (for 1% ≤MAF < 5%). Finally, intensity cluster plots of significant SNPs were visually inspected for over-dispersion biased no calling, and/or erroneous genotype assignment. SNPs exhibiting any of these characteristics were discarded.

Imputation For SNPs not directly genotyped, imputed data and proxy SNPs were used. Imputation was performed using the IMPUTE software package (v2) 5 using two reference panels, P0 (HapMap2, rel 22, combined CEU+YRI+ASN panels) and P1 (610k+, including the combined HumanHap610k and 1M reduced to 610k SNP content). Imputed SNPs were excluded if MAF<1%, Hardy-Weinberg p-value<10^{-6} and imputation quality score was <0.5. For proxy SNPs, genotype data was taken from the 1,000 Genomes Project database (http://www.1000genomes.org).

Statistical analysis

All statistical analyses were performed using the SNPassoc R package (version 2.10.0) (Gonzalez JR, 2007). Departure from Hardy–Weinberg equilibrium (HWE) was assessed in controls using the χ^2 test (1-degree of freedom, 1-df). Unconditional and conditional logistic regression analyses were used to assess associations. The primary association test assumed a log-additive genotypic effect, while the secondary analysis evaluated the co-dominant, dominant, and recessive effects following the Akaike criteria for the best genetic model. All the reported ORs and 95% CI were adjusted for age and gender. Bonferroni and false discovery rate (FDR) approaches were used to correct for multiple testing (Benjamini Y et al., 2001).

In the Leeds study, nevus number was treated as a dichotomous variable excluding individuals with 50-100 nevi and as a continuous variable, including all the individuals so as to apply the same criteria used in the BCN and TwinsUK sets, respectively. For the cases from Barcelona and Leeds, statistical heterogeneity was tested by two-sided Fisher's exact test.

In the TwinsUK cohort, statistical analysis for SNPs directly genotyped, imputed and proxies were performed using PLINK (http://pngu.mgh.harvard.edu/~purcell/plink/).

Imputation was performed using the IMPUTE software package (v2) 5 using two reference panels, P0 (HapMap2, rel 22, combined CEU+YRI+ASN panels) and P1 (610k+, including the combined HumanHap610k and 1M reduced to 610k SNP content).

The analysis of association between SNP and nevus number, continuous trait, was assessed in 3,054 individuals by regressing absolute nevus number and logarithm (log) nevus number on genotype under trend model. For SNPs directly genotyped and SNPs imputed, significant associations were considered when $P \leq 0.05$ in both statistical analyses, whereas for proxy SNPs only $P \leq 0.05$ for log nevus number analysis was considered. Log nevus number analysis was repeated accounting for family structure via the lme4 function in R (http://www.r-project.org). Age was included in all models but not gender information as all TwinsUK individuals were females.

To test independent association of SNP with nevus number, association between polymorphisms and confounding variables (gender, skin type and number melanoma) was assessed.

Conflicts of interest

The authors declare no conflicts of interest.

Acknowledgements

This study has been supported by Fondo de Investigaciones Sanitarias (FIS grants 05/0302, 06/0265 and 09/1393). Zighereda Ogbah was supported by financial support provided by Agència de Gestió d'Ajuts Universitaris i de Recerca (AGAUR) and by GenoMEL Exchange Programme fund by the European Commission under the 6th Framework Programme.

We also thank the principal investigators of the GenoMEL funding grants David Elder, Julia Newton Bishop and Nelleke Gruis for their support and to the leader of the exchange Programme Giovanna Bianchi-Scarrá and also to Linda Battistuzzi.

Thanks to research nurses from Melanoma Unit, Daniel Gabriel and Pablo Iglesias for collecting phenotypic data from patients; to Maria Magdalena Monfort from CeGen (Barcelona) and Michael Churchman from LIMM (Leeds) for genotyping support.

This work was developed in the Melanoma Units of Hospital Clinic I Provincial de Barcelona, IDIBAPS and in Epidemiology and Biostatistics, St James Hospital, LIMM, Leeds.

The MuTHER study was funded by the Wellcome Trust Program grant # 081917. Genotyping of TwinsUK samples was provided by the Wellcome Trust Sanger Institute and the National Eye Institute via a US National Institutes of Health (NIH)/Center for Inherited Disease Research (CIDR) genotyping project. TwinsUK was also supported in part by Wellcome Trust grant WT098051 to the Wellcome Trust Sanger Institute. TwinsUK also receives support from the ENGAGE project grant agreement HEALTH-F4-2007-201413 and from the Department of Health via the National Institute for Health Research (NIHR) comprehensive Biomedical Research Centre award to Guy's & St. Thomas' National Health Service Foundation Trust in partnership with King's College London

References

Aguilera, P., Puig, S., Guilabert, A., Julià, M., Romero, D., Vicente, A., González-Enseñat, M.A., Malvehy, J. (2009). Prevalence study of nevi in children from Barcelona. Dermatology 218, 203-14.

Andrew, T., Hart, D. J., Snieder, H., De Lange, M., Spector, T. D., and Macgregor, A. J. (2001). Are twins and singletons comparable? A study of disease-related and lifestyle characteristics in adult women. Twin Res 4, 464-77.

Bahuau, M., Vidaud, D., Kujas, M., Palangie, A., Assouline, B., Chaignaud-Lebreton, M., Prieur, M., Vidaud, M., Harpey, J. P., Lafourcade, J., et al. (1997). Familial aggregation of malignant melanoma/dysplastic naevi and tumours of the nervous system: an original syndrome of tumour proneness. Ann Genet 40, 78-91.

Barrett, J. H., Iles, M. M., Harland, M., Taylor, J. C., Aitken, J. F., Andresen, P. A., Akslen, L. A., Armstrong, B. K., Avril, M. F., Azizi, E., et al. (2011). Genome-wide association study identifies three new melanoma susceptibility loci. Nat Genet 43, 1108-13.

Bataille, V., Bishop, J. A., Sasieni, P., Swerdlow, A. J., Pinney, E., Griffiths, K., and Cuzick, J. (1996). Risk of cutaneous melanoma in relation to the numbers, types and sites of naevi: a case-control study. Br J Cancer 73, 1605-11.

Bataille, V., Snieder, H., Macgregor, A. J., Sasieni, P., and Spector, T. D. (2000). Genetics of risk factors for melanoma: an adult twin study of nevi and freckles. J Natl Cancer Inst 92, 457-63.

Benjamini, Y., Drai, D., Elmer, G., Kafkafi, N., and Golani, I. (2001). Controlling the false discovery rate in behavior genetics research. Behav Brain Res 125, 279-84.

Bishop, J. A., Wachsmuth, R. C., Harland, M., Bataille, V., Pinney, E., Mac, K. P., Baglietto, L., Cuzick, J., and Bishop, D. T. (2000). Genotype/phenotype and penetrance studies in melanoma families with germline CDKN2A mutations. J Invest Dermatol 114, 28-33.

Blake, J. A., and Ziman, M. R. (2005). Pax3 transcripts in melanoblast development. Dev Growth Differ 47, 627-35.

Bonvin, E., Falletta, P., Shaw, H., Delmas, V., and Goding, C.R. (2012). A phosphatidylinositol 3-kinase-Pax3 axis regulates Brn-2 expression in melanoma. Mol Cell Biol 32, 4674-83.

Briollais, L., Chompret, A., Guilloud-Bataille, M., Bressac-De Paillerets, B., Avril, M. F., and Demenais, F. (2000). Patterns of familial aggregation of three melanoma risk factors: great number of naevi, light phototype and high degree of sun exposure. Int J Epidemiol 29, 408-15.

Chalepakis, G., Jones, F. S., Edelman, G. M., and Gruss, P. (1994). Pax-3 contains domains for transcription activation and transcription inhibition. Proc Natl Acad Sci U S A 91, 12745-9.

Chang, Y. M., Newton-Bishop, J. A., Bishop, D. T., Armstrong, B. K., Bataille, V., Bergman, W., Berwick, M., Bracci, P. M., Elwood, J. M., Ernstoff, M. S., et al. (2009). A pooled analysis of melanocytic nevus phenotype and the risk of cutaneous melanoma at different latitudes. Int J Cancer 124, 420-8.

Corry, G.N., and Underhill, D.A. (2005). Pax3 target gene recognition occurs through distinct modes that are differentially affected by disease-associated mutations. Pigment Cell Res. 18, 427-38.

Duffy, D. L., Iles, M. M., Glass, D., Zhu, G., Barrett, J. H., Hoiom, V., Zhao, Z. Z., Sturm, R. A., Soranzo, N., Hammond, C., et al. (2010). IRF4 variants have age-specific effects on nevus count and predispose to melanoma. Am J Hum Genet 87, 6-16.

Easton, D. F., Cox, G. M., Macdonald, A. M., and Ponder, B. A. (1991). Genetic susceptibility to naevi--a twin study. Br J Cancer 64, 1164-7.

Falchi, M., Spector, T. D., Perks, U., Kato, B. S., and Bataille, V. (2006). Genome-wide search for nevus density shows linkage to two melanoma loci on chromosome 9 and identifies a new QTL on 5q31 in an adult twin cohort. Hum Mol Genet 15, 2975-9.

Falchi, M., Bataille, V., Hayward, N. K., Duffy, D. L., Bishop, J. A., Pastinen, T., Cervino, A., Zhao, Z. Z., Deloukas, P., Soranzo, N., et al. (2009). Genome-wide association study identifies variants at 9p21 and 22q13 associated with development of cutaneous nevi. Nat Genet 41, 915-9.

Fernandez, L.P., Milne, R.L., Pita, G., Floristan, U., Sendagorta, E., Feito, M., Avilés, J.A., Martin-Gonzalez, M., Lázaro, P., Benitez, J., and Ribas, G. (2009). Pigmentation-related genes and their implication in malignant melanoma susceptibility. Exp Dermatol. 18, 634-42.

Florell, S. R., Meyer, L. J., Boucher, K. M., Porter-Gill, P. A., Hart, M., Erickson, J., Cannon-Albright, L. A., Pershing, L. K., Harris, R. M., Samlowski, W. E., et al. (2004). Longitudinal assessment of the nevus phenotype in a melanoma kindred. J Invest Dermatol 123, 576-82.

Gandini, S., Sera, F., Cattaruzza, M. S., Pasquini, P., Abeni, D., Boyle, P., and Melchi, C. F. (2005). Meta-analysis of risk factors for cutaneous melanoma: I. Common and atypical naevi. Eur J Cancer 41, 28-44.

Gonzalez, JR., Armengol, L., Sole, X., Guino, E., Mercader, JM., Estivil,l X., Moreno, V. (2007). SNPassoc: an R package to perform whole genome association studies. Bioinformatics 23, 644-645.

He, S., Yoon, H. S., Suh, B. J., and Eccles, M. R. (2010). PAX3 Is extensively expressed in benign and malignant tissues of the melanocytic lineage in humans. J Invest Dermatol 130, 1465-8.

Koyanagi, K., Kuo, C., Nakagawa, T., Mori, T., Ueno, H., Lorico, A. R., Jr., Wang, H. J., Hseuh, E., O'day, S. J., and Hoon, D. S. (2005). Multimarker quantitative real-time PCR detection of circulating melanoma cells in peripheral blood: relation to disease stage in melanoma patients. Clin Chem 51, 981-8.

Kubic, J. D., Young, K. P., Plummer, R. S., Ludvik, A. E., and Lang, D. (2008). Pigmentation PAX-ways: the role of Pax3 in melanogenesis, melanocyte stem cell maintenance, and disease. Pigment Cell Melanoma Res 21, 627-45.

Lang, D., Lu, M. M., Huang, L., Engleka, K. A., Zhang, M., Chu, E. Y., Lipner, S., Skoultchi, A., Millar, S. E., and Epstein, J. A. (2005). Pax3 functions at a nodal point in melanocyte stem cell differentiation. Nature 433, 884-7.

Leiter, U., and Garbe, C. (2008). Epidemiology of melanoma and nonmelanoma skin cancer--the role of sunlight. Adv Exp Med Biol 624, 89-103.

Medic, S., and Ziman, M. (2009). PAX3 across the spectrum: from melanoblast to melanoma. Crit Rev Biochem Mol Biol 44, 85-97.

Medic, S., and Ziman, M. (2010). PAX3 expression in normal skin melanocytes and melanocytic lesions (naevi and melanomas). PLoS One 5, e9977.

Medic, S., Rizos, H., and Ziman, M. (2011). Differential PAX3 functions in normal skin melanocytes and melanoma cells. Biochem Biophys Res Commun 411, 832-7.

Nan, H., Xu, M, Zhang, J., Zhang, M., Kraft, P., Qureshi, AA., Chen, C., Guo, Q., Hu, FB., Rimm, EB., et al.(2011). Genome-wide association study identifies nidogen 1 (NID1) as a susceptibility locus to cutaneous nevi and melanoma risk. Hum Mol Genet 20, 2673-9.

Nica, A.C., Parts, L., Glass, D., Nisbet, J., Barrett, A., Sekowska, M., Travers, M., Potter, S., Grundberg, E., Small, K., et al. (2011) The architecture of gene regulatory variation across multiple human tissues: the MuTHER study**.** PLoS Genet **7:**e1002003.

Plummer, R. S., Shea, C. R., Nelson, M., Powell, S. K., Freeman, D. M., Dan, C. P., and Lang, D. (2008). PAX3 expression in primary melanomas and nevi. Mod Pathol 21, 525-30.

Pritchard, C., Grosveld, G., Hollenbach, A.D. (2003). Alternative splicing of Pax3 produces a transcriptionally inactive protein. Gene 305, 61-9.

Puig, S., Ruiz, A., Castel, T., Volpini, V., Malvehy, J., Cardellach, F., Lynch, M., Mascaro, J. M., and Estivill, X. (1997). Inherited susceptibility to several cancers but absence of linkage between dysplastic nevus syndrome and CDKN2A in a melanoma family with a mutation in the CDKN2A (P16INK4A) gene. Hum Genet 101, 359-64.

Randerson-Moor, J.A., Taylor, J.C., Elliott, F., Chang, Y.M., Beswick, S., Kukalizch, K., Affleck, P., Leake, S., Haynes, S., Karpavicius, B., et al. (2009). Vitamin D receptor gene polymorphisms, serum 25-hydroxyvitamin D levels, and melanoma: UK case-control comparisons and a meta-analysis of published VDR data. Eur J Cancer 45, 3271-81.

Sauna, Z. E., and Kimchi-Sarfaty, C. (2011). Understanding the contribution of synonymous mutations to human disease. Nat Rev Genet 12, 683-91.

Scholl, F. A., Kamarashev, J., Murmann, O. V., Geertsen, R., Dummer, R., and Schafer, B. W. (2001). PAX3 is expressed in human melanomas and contributes to tumor cell survival. Cancer Res 61, 823-6.

Takeuchi, H., Morton, D. L., Kuo, C., Turner, R. R., Elashoff, D., Elashoff, R., Taback, B., Fujimoto, A., and Hoon, D. S. (2004). Prognostic significance of molecular upstaging of paraffin-embedded sentinel lymph nodes in melanoma patients. J Clin Oncol 22, 2671-80.

Wachsmuth, R. C., Gaut, R. M., Barrett, J. H., Saunders, C. L., Randerson-Moor, J. A., Eldridge, A., Martin, N. G., Bishop, T. D., and Newton Bishop, J. A. (2001). Heritability and gene-environment interactions for melanocytic nevus density examined in a U.K. adolescent twin study. J Invest Dermatol 117, 348-52.

Wachsmuth, R. C., Turner, F., Barrett, J. H., Gaut, R., Randerson-Moor, J. A., Bishop, D. T., and Bishop, J. A. (2005). The effect of sun exposure in determining nevus density in UK adolescent twins. J Invest Dermatol 124, 56-62.

Yang, G., Li, Y., Nishimura, E. K., Xin, H., Zhou, A., Guo, Y., Dong, L., Denning, M. F., Nickoloff, B. J., and Cui, R. (2008). Inhibition of PAX3 by TGF-beta modulates melanocyte viability. Mol Cell 32, 554-63.

Yang, T.P., Beazley, C., Montgomery, S.B., Dimas, A.S., Gutierrez-Arcelus, M., Stranger, B.E., Deloukas, P., Dermitzakis, E.T. (2010). Genevar: a database and Java application for the analysis and visualization of SNP-gene associations in eQTL studies**.** Bioinformatics 26**,** 2474-76

Yokoyama, S., Woods, S. L., Boyle, G. M., Aoude, L. G., Macgregor, S., Zismann, V., Gartside, M., Cust, A. E., Haq, R., Harland, M., et al. (2011). A novel recurrent mutation in MITF predisposes to familial and sporadic melanoma. Nature 480, 99-103.

Zhu, G., Duffy, D. L., Eldridge, A., Grace, M., Mayne, C., O'gorman, L., Aitken, J. F., Neale, M. C., Hayward, N. K., Green, A. C., et al. (1999). A major quantitative-trait locus for mole density is linked to the familial melanoma gene CDKN2A: a maximum-likelihood combined linkage and association analysis in twins and their sibs. Am J Hum Genet 65, 483-92.

Zhu, G., Montgomery, G. W., James, M. R., Trent, J. M., Hayward, N. K., Martin, N. G., and Duffy, D. L. (2007). A genome-wide scan for naevus count: linkage to CDKN2A and to other chromosome regions. Eur J Hum Genet 15, 94-102.

Supplementary Figure Legends

Figure 1

Schematic illustration of genomic structure of *PAX3* and location of genotyped SNPs.

Figure 2

A rs6754024

B rs6436309

C rs1549773

D rs13410020

A, B C and D plots were generated using the Genvar web tool (SNP-gene association analysis) and published expression data from Nica, et al., 2011. The results reported are from eQTL data of Twins dataset of MuTHER study. The Twins pairs were separated in two unrelated groups. The individuals are plotted on stripcharts. The four panels showed the results of the correlation between *PAX3* SNP and *PAX3* transcript levels in three different tissues, examined using Spearman. The observed and permutation *P* values are shown for each analysis. The four *PAX3* SNPs minor alleles were correlated with decreased *PAX3* gene expression. The SNP rs6754024, the only genotyped, has been associated with high nevus number in melanoma cases from BCN and Leeds. The SNP rs6436309 is in linkage (r2=0.53) with rs6754024. The SNPs rs1549773 (r2=1) and rs13410020 (r2=0.68) are in LD with rs2855268; this SNP has been found to protect from high nevus number in BCN melanoma cases and Twins UK.

Abbreviations: A= Adipose tissue biopsy (n 160). L= Lymphoblastoid cell lines (n 166). S= skin tissue biopsy (n 160). eQTL= expression quantitative trait loci. Twin1 = unrelated twin group 1. Twin2 =unrelated twin group 2. rho = Spearman's rank correlation coefficient between genotype and transcript levels. P = p-value for correlation. Pemp = adjusted p-value by permutation test.

Table 1. ASSOCIATION BETWEEN *PAX3* SNPs AND NEVUS NUMBER IN MELANOMA CASES FROM BCN AND LEEDS

PAX3 SNP	Population	Genotype	Cases <50(%)	Cases >100(%)	OR (95%CI)	$P_{Additive*}$
rs6754024	BCN	G/G	112 (78.3)	70 (65.4)	1.00	
		G/T	29 (20.3)	32 (29.9)	1.92 (1.05- 3.53)	
		T/T	2 (1.4)	5 (4.7)	4.71 (0.87- 25.53)	
		T			2.00 (1.20- 3.34)	0.0066
	LEEDS	G/G	383 (70.1)	129 (70.9)	1.00	
		G/T	151 (27.7)	42 (23.1)	0.83 (0.55- 1.24)	
		T/T	12 (2.2)	11 (6.0)	3.39 (1.41- 8.15)	
		T			1.16 (0.84- 1.60)	0.366
		T/T			3.57 (1.50- 8.52)	0.0053 $P_{recessive}$
rs7600206	BCN	T/T	74 (54.0)	77 (75.5)	1.00	
		C/T	57 (41.6)	23 (22.5)	0.33 (0.18- 0.61)	
		C/C	6 (4.4)	2 (2.0)	0.32 (0.06- 1.72)	
		C			0.38 (0.22-0.66)	0.0002
	LEEDS	T/T	373 (68.2)	116 (63.7)	1.00	
		C/T	155 (28.3)	60 (33.0)	1.24 (0.85- 1.80)	
		C/C	19 (3.5)	6 (3.3)	0.91 (0.35- 2.40)	
		C			1.12 (0.83- 1.53)	0.649
	BCN	T/T	48 (35.8)	58 (56.9)	1.00	
		C/T	74 (55.2)	40 (39.2)	0.43 (0.25- 0.76)	
		C/C	12 (9.0)	4 (3.9)	0.25 (0.07- 0.86)	
		C			0.46 (0.29- 0.74)	0.0008

rs12995399						
	LEEDS	T/T	336 (61.5)	99 (54.4)	1.00	
		C/T	184 (33.7)	74 (40.7)	1.28 (0.89- 1.83)	
		C/C	26 (4.8)	9 (4.9)	1.12 (0.50- 2.53)	
		C			1.17 (0.88- 1.57)	0.280
	BCN	T/T	103 (73.0)	53 (52.5)	1.00	
		C/T	37 (26.2)	44 (43.6)	2.41 (1.36- 4.28)	
		C/C	1 (0.7)	4 (4.0)	10.04 (1.01- 9.90)	
rs10180903		C			2.54 (1.51- 4.29)	0.0003
	LEEDS	T/T	357 (65.3)	113 (62.4)	1.00	
		C/T	167 (30.5)	59 (32.6)	1.05 (0.72- 1.53)	
		C/C	23 (4.2)	9 (5.0)	1.39 (0.60- 3.21)	
		C			1.10 (0.82- 1.49)	0.521
	BCN	A/A	56 (38.1)	20 (18.5)	1.00	
		A/C	66 (44.9)	54 (50.0)	2.22 (1.18- 4.20)	
		C/C	25 (17.0)	34 (31.5)	4.06 (1.93- 8.55)	
rs1013262		C			2.02 (1.39- 2.93)	0.0001
	LEEDS	A/A	185 (33.8)	56 (30.8)	1.00	
		A/C	263 (48.1)	97 (53.3)	1.22 (0.83- 1.81)	
		C/C	99 (18.1)	29 (15.9)	0.98 (0.58- 1.67)	
		C			1.03 (0.80- 1.32)	0.837
	BCN	G/G	103 (70.5)	92 (85.2)	1.00	
		C/G	39 (26.7)	15 (13.9)	0.44 (0.23- 0.87)	
		C/C	4 (2.7)	1 (0.9)	0.29 (0.03- 2.71)	
rs2855268		C			0.46 (0.26- 0.84)	0.008
	LEEDS	G/G	403 (73.7)	135 (74.2)	1.00	
		C/G	137 (25.0)	43 (23.6)	0.88 (0.59- 1.32)	
		C/C	7 (1.3)	4 (2.2)	2.28 (0.60- 8.62)	
		C			1.00 (0.69- 1.43)	0.984
	BCN	C/C	51 (37.5)	54 (54.0)	1.00	
rs1978859**		C/T	66 (48.5)	45 (45.0)	0.65 (0.37- 1.12)	
		T/T	19 (14.0)	1 (1.0)	0.06 (0.01- 0.45)	
		T			0.47 (0.30- 0.75)	0.0009

Cases <50= cases with less than 50 nevi; Cases >100= case with greater than 100 nevi. *Adjusted for age and gender **SNP rs1978859 failed genotypyng in Leeds samples.

Table 2. *PAX3* SNP rs6754024 ASSOCIATION WITH NEVUS NUMBER IN MELANOMA CASES FROM BCN AND LEEDS

PAX3 rs6754024	MAF T allele	Genotype	Cases <50(%)	Cases >100(%)	$OR_{Additive}$(95%CI)	$OR_{Recessive}$(95%CI)	$P_{Additive*}$	$P_{Recessive*}$
		G/G	112 (78.3)	70 (65.4)				
		G/T	29 (20.3)	32 (29.9)				
BCN cases	0.15	T/T	2 (1.4)	5 (4.7)	**2.00 (1.20- 3.34)**	4.71 (0.87- 25.53)	**0.0066**	0.089
		G/G	383 (70.1)	129 (70.9)				
		G/T	151 (27.7)	42 (23.1)				
LEEDS cases	0.16	T/T	12 (2.2)	11 (6.0)	1.16 (0.84- 1.60)	**3.57 (1.50- 8.52)**	0.366	**0.0053**
BCN +		G/G	495 (71.8)	199 (68.8)				
LEEDS**		G/T	180 (26.2)	74 (25.7)				
cases	0.16	T/T	14 (2.0)	16 (5.6)	**1.39 (1.06- 1.82)**	**3.96 (1.78- 8.78)**	**0.0135**	**0.0009**

Significant results are in bold. MAF= minor allele frequency; Cases <50= cases with less than 50 nevi; Cases >100= case with greater than 100 nevi. *Adjusted for age, gender and **center

Table 3. ASSOCIATIONS BETWEEN *PAX3* SNPs AND NEVUS NUMBER IN HEALTHY TWINSUK

a) Association between PAX3 SNPs and nevus number

				TOTAL NEVUS NUMBER			LOG NEVUS NUMBER			LOG NEVUS NUMBER [1]		
	SNP	**BP**	**A1**	**BETA**	***P***	**N**	**BETA**	***P***	**N**	**BETA**	***P***	**N**
1	**rs7600206**	**222864831**	**C**	**5.303**	**0.0001**	**2725**	**0.15**	**0.0005**	**2725**	**0.13423**	**0.0020**	**2704**
2	**rs12995399**	**222845245**	**C**	**4.16**	**0.0005**	**3012**	**0.1108**	**0.0032**	**3012**	**0.09044**	**0.0163**	**2991**
3	**rs2855268**	**222775163**	**G***	**-2.78**	**0.0503**	**3006**	**-0.1283**	**0.0038**	**3006**	**-0.1263**	**0.0057**	**2985**
4	rs1013262	222799023	G*	**-2.438**	**0.0179**	**2901**	-0.03069	0.347	2901	-0.02523	0.448	2880
5	rs10180903	222823973	C	-2.143	0.156	2648	-0.00603	0.899	2648	-0.01923	0.686	2628
6	rs1978859	222790575	T	-0.8186	0.439	3054	-0.04473	0.175	3054	-0.0353	0.293	3033
7	rs6754024	222870268	T	-0.2347	0.863	3054	-0.01569	0.712	3054	-0.03055	0.476	3033

b) Association between proxy PAX3 SNPS and nevus number

	SNP	**BP**	**A1**	**BETA**	***P***	**N**
1	**rs10932952**	**222853916**	**C**	**3.819**	**0.0013**	**3053**
2	**rs7588554**	**222782895**	**G**	**-2.255**	**0.0276**	**3054**
3	rs13385121	222710949	G	-2.81	0.059	3054
4	rs1549773	222804212	T	-2.446	0.076	3054

c) LD analysis

	SNP	**Proxy**	**Distance**	r^2	**D'**	**Coordinate_HG18**
1	rs12995399	rs10932952	8671	0.945	1	222853916
2	rs1013262	rs7588554	16128	0.966	1	222782895

Significant results are in bold.

a) Results of total nevus number and logarithm (log) nevus number analysis, both age –adjusted, for the 7 PAX3 SNPs found previously associated with nevus number in Barcelona study.
[1]Log nevus number analysis results, accounting for family structure and age –adjusted.
*Minor allele for SNPs rs1013262 (G) and rs2855268 (G) differs from that presented in Table 1 for BCN and Leeds studies as in TwinsUK the opposite strand was studied for both SNPs.

b) Log nevus number analysis results for proxy PAX3 SNPs.

c) LD analysis showed association degree between PAX3 SNPs rs12995399 and rs1013262 and proxy SNPs rs10932952 and rs7588554, respectively.

Abbreviations
BP= genomic position in base-pair; *P*=P values are from trend model; A1= minor allele. BETA= Effect size (beta) is expressed for nevus number and log scale nevus number. N= number of individuals genotyped. r^2=r-squared. D'= DPrime. LD=linkage disequilibrium

Table 4. GENERAL INFORMATION OF THE *PAX3* SNPs ANALYZED IN BCN, LEEDS AND TWINSUK COHORTS

SNP	GENOMIC POSITION	FUNCTION	GENOTYPE	MINOR ALELE	MAF BCN MM cases	MAF LEEDS MM cases	MAF TWINSUK Healthy individuals
rs6754024	222870268	INTRON 1	G/T	T	0.15	0.16	0.15
rs7600206	222864831	INTRON 4	C/T	C	0.20	0.17	0.17
rs12995399	222845245	INTRON 4	C/T	C	0.31	0.22	0.22
rs10180903	222823973	INTRON 4	C/T	C	0.19	0.20	0.15
rs1013262	222799023	INTRON 6	A/C	C*	0.46	0.42	0.40
rs1978859	222790575	INTRON 7	C/T	C	0.32	-**	0.36
rs2855268	222775163	INTRON 7	C/G	C*	0.13	0.15	0.15

*In TwinsUK, the minor allele for SNPs rs1013262 and rs2855268 was (G) as the opposite strand was studied. **SNP rs1978859 failed genotyping in Leeds samples. MAF= minor allele frequency, MM cases= melanoma cases.